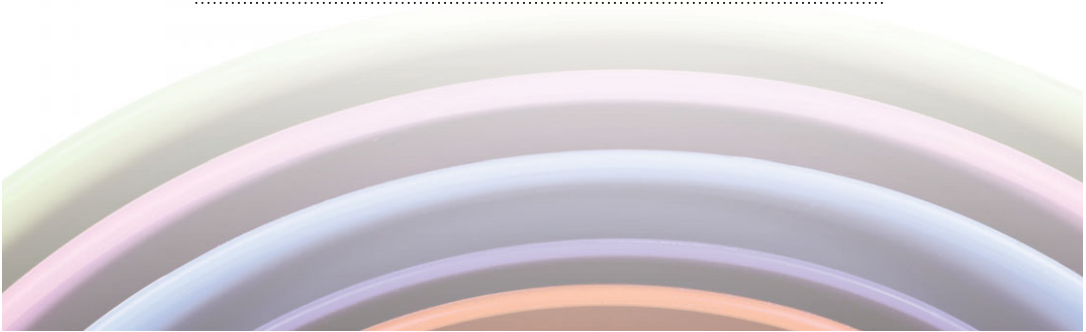


8

Sampling

Chapter outline

Introduction to survey research	184
Introduction to sampling	186
Sampling error	188
Types of probability sample	190
Simple random sample	190
Systematic sample	191
Stratified random sampling	192
Multi-stage cluster sampling	193
The qualities of a probability sample	195
Sample size	197
Absolute and relative sample size	197
Time and cost	198
Non-response	199
Heterogeneity of the population	200
Kind of analysis	201
Types of non-probability sampling	201
Convenience sampling	201
Snowball sampling	202
Quota sampling	203
Limits to generalization	205
Error in survey research	205
<i>Key points</i>	206
<i>Questions for review</i>	206





Chapter guide

This chapter and the three that follow it are very much concerned with principles and practices associated with social survey research. Sampling principles are not exclusively concerned with survey research; for example, they are relevant to the selection of documents for content analysis (see Chapter 13). However, in this chapter the emphasis will be on sampling in connection with the selection of people who would be asked questions by interview or questionnaire. The chapter explores:

- the role of sampling in relation to the overall process of doing survey research;
- the related ideas of generalization (also known as external validity) and of a representative sample; the latter allows the researcher to generalize findings from a sample to a population;
- the idea of a *probability sample*—that is, one in which a random selection process has been employed;
- the main types of probability sample: the simple random sample; the systematic sample; the stratified random sample; and the multi-stage cluster sample;
- the main issues involved in deciding on sample size;
- different types of non-probability sample, including quota sampling, which is widely used in market research and opinion polls;
- potential sources of error in survey research.

Introduction to survey research

This chapter is concerned with some important aspects of conducting a survey, but it presents only a partial picture, because there are many other steps. In this chapter we are concerned with the issues involved in selecting individuals for survey research, although the principles involved apply equally to other approaches to quantitative research, such as content analysis. Chapters 9, 10, and 11 deal with the data-collection aspects of conducting a survey, while Chapters 15 and 16 deal with issues to do with the analysis of data.

Figure 8.1 aims to outline the main steps involved in doing survey research. Initially, the survey will begin with general research issues that need to be investigated. These are gradually narrowed down so that they become research questions, which may take the form of hypotheses, but this need not necessarily be the case. The movement from research issues to research questions is likely to be the result of reading the literature relating to the issues, such as relevant theories and evidence (see Chapters 1 and 4).

Once the research questions have been formulated, the planning of the fieldwork can begin. In practice, decisions relating to sampling and the research instrument will overlap, but they are presented in Figure 8.1 as part of a sequence. The figure is meant to illustrate the main phases of a survey, and these different steps (other than those to do with sampling, which will be covered in this chapter) will be followed through in Chapters 9–11 and 15–16.

The survey researcher needs to decide what kind of population is suited to the investigation of the topic and also needs to formulate a research instrument and how it should be administered. By ‘research instrument’ is meant simply something like a **structured interview** schedule or a **self-completion questionnaire**. Moreover, there are several different ways of administering such instruments. Figure 8.2 outlines the main types that are likely to be encountered. Types 1 through 4 are covered in Chapter 9. Types 5 and 6 are covered in Chapter 10. Types 7 through 9 are covered in Chapter 28 in the context of the use of the Internet generally.

Figure 8.1

Steps in conducting a social survey

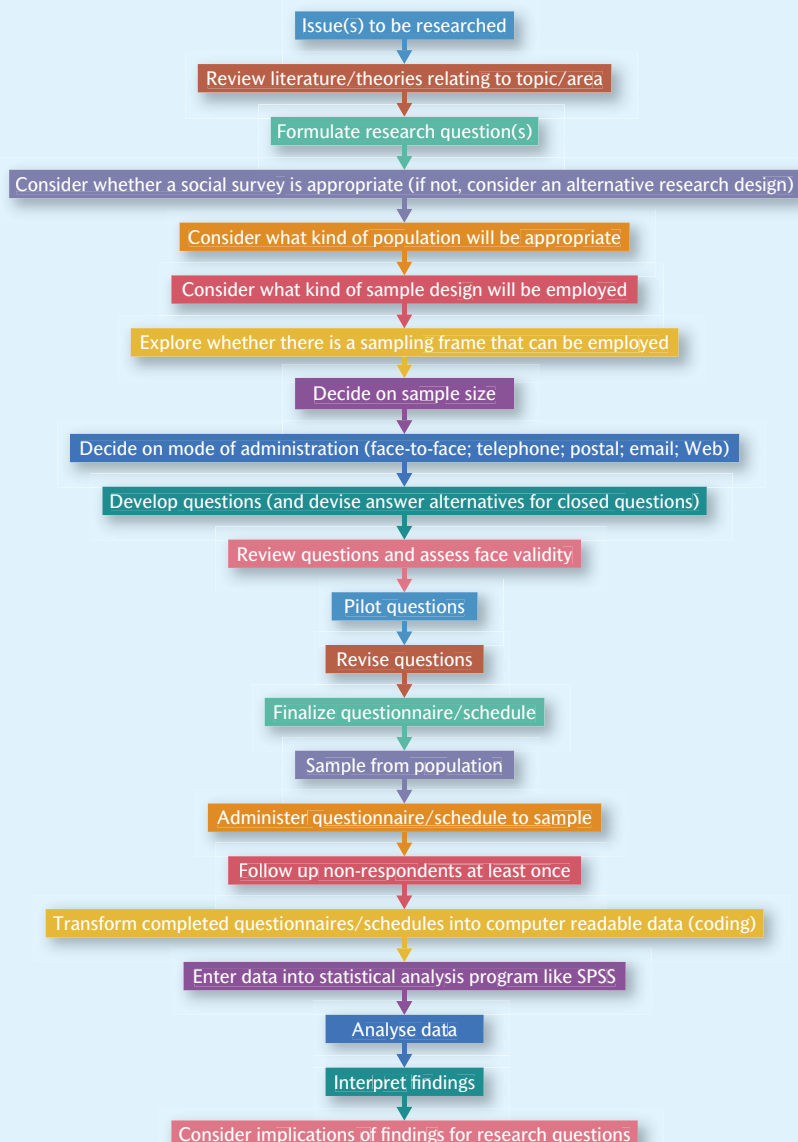
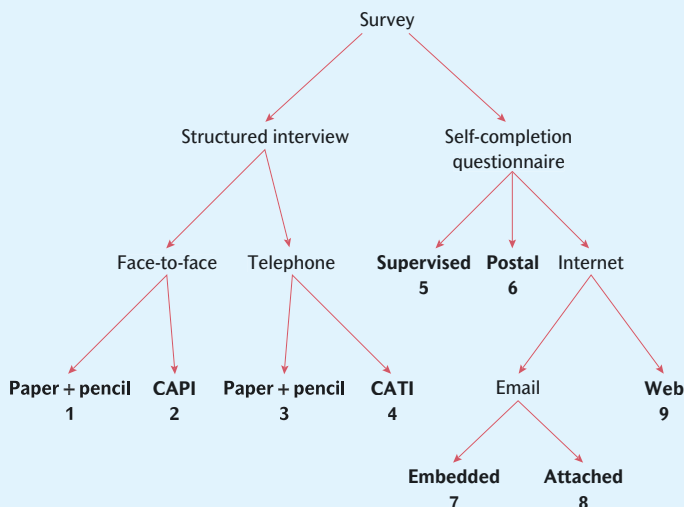


Figure 8.2

Main modes of administration of a survey



Notes: CAPI is computer-assisted personal interviewing; CATI is computer-assisted telephone interviewing.



Introduction to sampling

Many of the readers of this book will be university or college students. At some point in your stay at your university (I will use this term from now on to include colleges) you may have wondered about the attitudes of your fellow students to various matters, or about their behaviour in certain areas, or something about their backgrounds. If you were to decide to examine any or all of these three areas, you might consider conducting structured interviews or sending out questionnaires in order to find out about their behaviour, attitudes, and backgrounds. You will, of course, have to consider how best to design your interviews or questionnaires, and the issues that are involved in the decisions that need to be made about designing these research instruments and administering them will be the focus of Chapters 9–11. However, before getting to that point you are likely to be confronted with a problem. Let us say that your university is quite large and has around 9,000 students. It is extremely unlikely that you will have the time and

resources to conduct a survey of all these students. It is unlikely that you would be able to send questionnaires to all 9,000 and even more unlikely that you would be able to interview all of them, since conducting survey research by interview is considerably more expensive and time consuming, all things being equal, than by postal questionnaire (see Chapter 10). It is almost certain that you will need to *sample* students from the total population of students in your university.

The need to sample is one that is almost invariably encountered in quantitative research. In this chapter I will be almost entirely concerned with matters relating to sampling in relation to social survey research involving data collection by structured interview or questionnaire. Other methods of quantitative research involve sampling considerations, as will be seen in Chapters 12 and 13, when we will examine structured observation and content analysis respectively. The principles of sampling involved are more or less identical in connection with

these other methods, but frequently other considerations come to the fore as well.

But will any old sample suffice? Would it be sufficient to locate yourself in a central position on your campus (if it has one) and then interview the students who come past you and whom you are in a position to interview? Alternatively, would it be sufficient to go around your student union asking people to be interviewed? Or again to send questionnaires to everyone on your course?

The answer, of course, depends on whether you want to be able to *generalize* your findings to the entire student body in your university. If you do, it is unlikely that any of the three sampling strategies proposed in the previous paragraph would provide you with a *representative sample* of all students in your university. In order to be able to generalize your findings from your sample to the population from which it was selected, the sample must be representative. See Key concept 8.1 for an explanation of key terms concerning sampling.



Key concept 8.1

Basic terms and concepts in sampling

- **Population:** basically, the universe of units from which the sample is to be selected. The term 'units' is employed because it is not necessarily people who are being sampled—the researcher may want to sample from a universe of nations, cities, regions, firms, etc. Finch and Hayes (1994), for example, based part of their research upon a random sample of wills. Their population, therefore, was a population of wills. Thus, 'population' has a much broader meaning than the everyday use of the term, whereby it tends to be associated with a nation's entire population.
- **Sample:** the segment of the population that is selected for investigation. It is a subset of the population. The method of selection may be based on a probability or a non-probability approach (see below).
- **Sampling frame:** the listing of all units in the population from which the sample will be selected.
- **Representative sample:** a sample that reflects the population accurately so that it is a microcosm of the population.
- **Sampling bias:** a distortion in the representativeness of the sample that arises when some members of the population (or more precisely the sampling frame) stand little or no chance of being selected for inclusion in the sample.
- **Probability sample:** a sample that has been selected using random selection so that each unit in the population has a known chance of being selected. It is generally assumed that a *representative sample* is more likely to be the outcome when this method of selection from the population is employed. The aim of probability sampling is to keep *sampling error* (see below) to a minimum.
- **Non-probability sample:** a sample that has not been selected using a random selection method. Essentially, this implies that some units in the population are more likely to be selected than others.
- **Sampling error:** error in the findings deriving from research due to the difference between a sample and the population from which it is selected. This may occur even though probability sampling has been employed.
- **Non-sampling error:** error in the findings deriving from research due to the differences between the population and the sample that arise either from deficiencies in the sampling approach, such as an inadequate sampling frame or **non-response** (see below), or from such problems as poor question wording, poor interviewing, or flawed processing of data.
- **Non-response:** a source of non-sampling error that is particularly likely to happen when individuals are being sampled. It occurs whenever some members of the sample refuse to cooperate, cannot be contacted, or for some reason cannot supply the required data (for example, because of mental incapacity).
- **Census:** the enumeration of an entire population. Thus, if data are collected in relation to all units in a population, rather than in relation to a sample of units of that population, the data are treated as census data. The phrase '*the census*' typically refers to the complete enumeration of all members of the population of a nation state—that is, a national census. This form of enumeration currently occurs once every ten years in the UK, although there is some uncertainty at the time of writing about whether another census will take place. However, in a statistical context, like the term *population*, the idea of a census has a broader meaning than this.

Why might the strategies for sampling students previously outlined be unlikely to produce a representative sample? There are various reasons, of which the following stand out.

- The first two approaches depend heavily upon the availability of students during the time or times that you search them out. Not all students are likely to be equally available at that time, so the sample will not reflect these students.
- They also depend on the students going to the locations. Not all students will necessarily pass the point where you locate yourself or go to the student union, or they may vary hugely in the frequency with which they do so. Their movements are likely to reflect such things as where their halls of residence or accommodation are situated, or where their departments are located, or their social habits. Again, to rely on these locations would mean missing out on students who do not frequent them.
- It is possible, not to say likely, that your decisions about which people to approach will be influenced by your judgements about how friendly or cooperative the people concerned are likely to be or by how comfortable you feel about interviewing students of the same (or opposite) gender to yourself, as well as by many other factors.
- The problem with the third strategy is that students on your course by definition take the same subject as each other and therefore will not be representative of all students in the university.

In other words, in the case of all of the three sampling approaches, your decisions about whom to sample are influenced too much by personal judgements, by prospective respondents' availability, or by your implicit criteria for inclusion. Such limitations mean that, in the language of survey sampling, your sample will be *biased*. A biased sample is one that does not represent the population from which the sample was selected. Sampling bias will occur if some members of the population

have little or no chance of being selected for inclusion in the sample. As far as possible, bias should be removed from the selection of your sample. In fact, it is incredibly difficult to remove bias altogether and to derive a truly representative sample. What needs to be done is to ensure that steps are taken to keep bias to an absolute minimum.

Three sources of sampling bias can be identified (see Key concept 8.1 for an explanation of key terms).

1. *If a non-probability or non-random sampling method is used.* If the method used to select the sample is not random, there is a possibility that human judgement will affect the selection process, making some members of the population more likely to be selected than others. This source of bias can be eliminated through the use of probability/random sampling, the procedure for which is described below.
2. *If the sampling frame is inadequate.* If the sampling frame is not comprehensive or is inaccurate or suffers from some other kind of similar deficiency, the sample that is derived cannot represent the population, even if a random/probability sampling method is employed.
3. *If some sample members refuse to participate or cannot be contacted—in other words, if there is non-response.* The problem with non-response is that those who agree to participate may differ in various ways from those who do not agree to participate. Some of the differences may be significant to the research question or questions. If the data are available, it may be possible to check how far, when there is non-response, the resulting sample differs from the population. It is often possible to do this in terms of characteristics such as gender or age, or, in the case of something like a sample of university students, whether the sample's characteristics reflect the entire sample in terms of faculty membership. However, it is usually impossible to determine whether differences exist between the population and the sample after non-response in terms of 'deeper' factors, such as attitudes or patterns of behaviour.



Sampling error

In order to appreciate the significance of sampling error for achieving a representative sample, consider Figures 8.3–8.7. Imagine we have a population of 200 people

and we want a sample of 50. Imagine as well that one of the variables of concern to us is whether people watch soap operas and that the population is equally divided

Figure 8.3

Watching soap operas in a population of 200

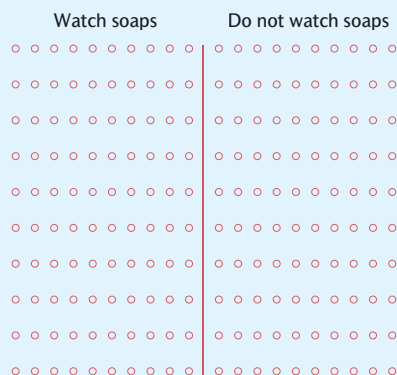


Figure 8.5

A sample with very little sampling error

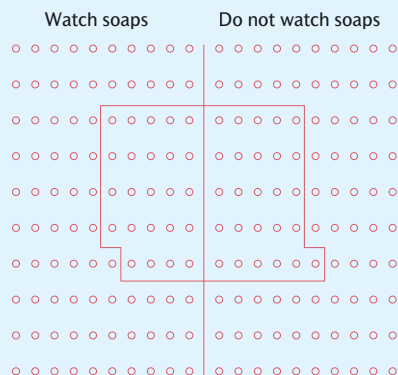


Figure 8.4

A sample with no sampling error

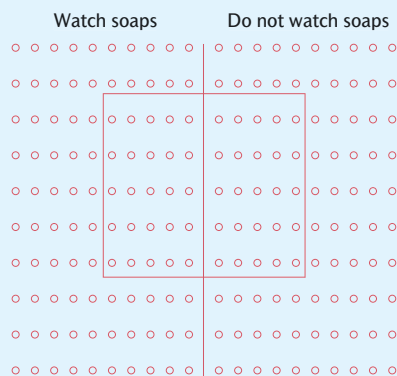
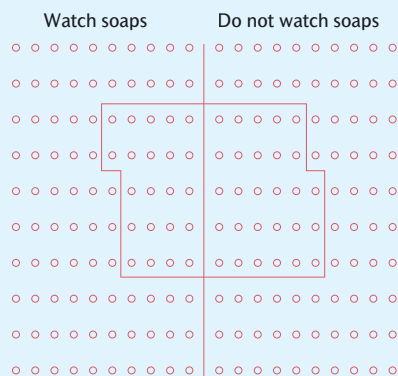


Figure 8.6

A sample with some sampling error

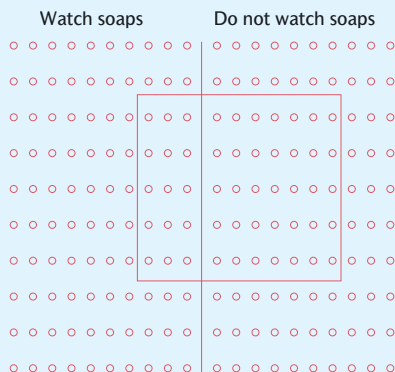


between those who do and those who do not. This split is represented by the vertical line that divides the population into two halves (Figure 8.3). If the sample is representative we would expect our sample of 50 to be equally split in terms of this variable (Figure 8.4). If there

is a small amount of sampling error, so that we have one person too many who does not watch soap operas and one too few who does, it will look like Figure 8.5. In Figure 8.6 we see a rather more serious degree of overrepresentation of people who do not watch soaps. This

Figure 8.7

A sample with a lot of sampling error



time there are three too many who do not watch them and three too few who do. In Figure 8.7 we have a very serious over-representation of people who do not watch soaps, because there are 35 people in the sample who do not watch them, which is much larger than the 25 who should be in the sample.

It is important to appreciate that, as suggested above, probability sampling does not and cannot eliminate sampling error. Even with a well-crafted probability sample, a degree of sampling error is likely to creep in. However, probability sampling stands a better chance than non-probability sampling of keeping sampling error in check so that it does not end up looking like the outcome featured in Figure 8.7. Moreover, probability sampling allows the researcher to employ tests of statistical significance that permit inferences to be made about the sample from which the sample was selected. These will be addressed in Chapter 15.



Types of probability sample

Imagine that we are interested in levels of alcohol consumption among university students and the variables that relate to variation in levels of drinking. We might decide to conduct our research in a single nearby university. This means that our population will be all students in that university, which will in turn mean that we will be able to generalize our findings only to students of that university. We simply cannot assume that levels of alcohol consumption and their correlates will be the same in other universities. We might decide that we want our research to be conducted only on full-time students, so that part-time students are omitted. Imagine too that there are 9,000 full-time students in the university.

Simple random sample

The **simple random sample** is the most basic form of probability sample. With random sampling, each unit of the population has an equal probability of inclusion in the sample. Imagine that we decide that we have enough money to interview 450 students at the university. This means that the probability of inclusion in the sample is

$$\frac{450}{9,000}, \text{ i.e. } 1 \text{ in } 20$$

This is known as the *sampling fraction* and is expressed as

$$\frac{n}{N}$$

where n is the sample size and N is the population size.

The key steps in devising our simple random sample can be represented as follows.

1. Define the population. We have decided that this will be all full-time students at the university. This is our N and in this case is 9,000.
2. Select or devise a comprehensive sampling frame. It is likely that the university will have an office that keeps records of all students and that this will enable us to exclude those who do not meet our criteria for inclusion—i.e. part-time students.
3. Decide your sample size (n). We have decided that this will be 450.

4. List all the students in the population and assign them consecutive numbers from 1 to N . In our case, this will be 1 to 9,000.
5. Using a table of random numbers, or a computer program that can generate random numbers, select n (450) different random numbers that lie between 1 and N (9,000).
6. The students to which the n (450) random numbers refer constitute the sample.

Two points are striking about this process. First, there is almost no opportunity for human bias to manifest itself. Students would not be selected on such subjective criteria as whether they looked friendly and approachable. The selection of whom to interview is entirely mechanical. Second, the process is not dependent on the students' availability. They do not have to be walking in the interviewer's proximity to be included in the sample. The process of selection is done without their knowledge. It is not until they are contacted by an interviewer that they know that they are part of a social survey.

Step 5 mentions the possible use of a table of random numbers. These can be found in the appendices of many statistics books. The tables are made up of columns of five-digit numbers, such as:

09188
90045
73189
75768
54016
08358
28306
53840
91757
89415

The first thing to notice is that, since these are five-digit numbers and the maximum number that we can sample from is 9,000, which is a four-digit number, none of the random numbers seems appropriate, except for 09188 and 08358, although the former is larger than the largest possible number. The answer is that we should take just four digits in each number. Let us take the last four digits. This would yield the following:

9188
0045
3189
5768
4016
8358
8306
3840
1757
9415

However, two of the resulting numbers—9188 and 9415—exceed 9,000. We cannot have a student with either of these numbers assigned to him or her. The solution is simple: we ignore these numbers. This means that the student who has been assigned the number 45 will be the first to be included in the sample; the student who has been assigned the number 3189 will be next; the student who has been assigned the number 5768 will be next; and so on.

However, this somewhat tortuous procedure may be replaced in some circumstances by using a *systematic sampling* procedure (see next section) and more generally can be replaced by enlisting the computer for assistance (see Tips and skills 'Generating random numbers').

Systematic sample

A variation on the simple random sample is the **systematic sample**. With this kind of sample, you select units directly from the sampling frame—that is, without resorting to a table of random numbers.

We know that we are to select 1 student in 20. With a systematic sample, we would make a random start between 1 and 20 inclusive, possibly by using the last two digits in a table of random numbers. If we did this with the ten random numbers above, the first relevant one would be 54016, since it is the first one where the last two digits yield a number of 20 or below, in this case 16. This means that the sixteenth student on our sampling frame is the first to be in our sample. Thereafter, we take every twentieth student on the list. So the sequence will go:

16, 36, 56, 76, 96, 116, etc.



Tips and skills

Generating random numbers

The method for generating random numbers described in the text is what might be thought of as the classic approach. However, a far neater and quicker way is to generate random numbers on the computer. For example, the following website provides an online random generator which is very easy to use:

www.psychicscience.org/random.aspx (accessed 9 August 2010).

If we want to select 450 cases from a population of 9,000, specify 450 after Generate, the digit 1 after random integers between and then 9000 after and. You will also need to specify from a drop-down menu 'with no repeats'. This means that no random number will be selected more than once. Then simply click on GO and the 450 random numbers will appear in a box below OUTPUT. You can then copy and paste the random numbers into a document.

This approach obviates the need to assign numbers to students' names and then to look up names of the students whose numbers have been drawn by the random selection process. It is important to ensure, however, that there is no inherent ordering of the sampling frame, since this may bias the resulting sample. If there is some ordering to the list, the best solution is to rearrange it.

Stratified random sampling

In our imaginary study of university students, one of the features that we might want our sample to exhibit is a proportional representation of the different faculties to which students are attached. It might be that the kind of discipline a student is studying is viewed as relevant to a wide range of attitudinal features that are relevant to the study of drinking. Generating a simple random sample or a systematic sample *might* yield such a representation, so that the proportion of humanities students in the sample is the same as that in the student population and so on. Thus, if there are 1,800 students in the humanities faculty, using our sampling fraction of 1 in 20, we would expect to have 90 students in our sample from this faculty. However, because of sampling error, it is unlikely that this will occur and that there will be a difference, so that there may be, say, 85 or 93 from this faculty.

Because it is almost certain that the university will include in its records the faculty in which students are based, or indeed may have separate sampling frames for each faculty, it will be possible to ensure that students are accurately represented in terms of their faculty membership. In the language of sampling, this means stratifying the population by a criterion (in this case, faculty membership) and selecting either a simple random sample or a systematic sample from each of the resulting strata. In

Table 8.1

The advantages of stratified sampling

Faculty	Population	Stratified sample	Hypothetical simple random or systematic sample
Humanities	1,800	90	85
Social sciences	1,200	60	70
Pure sciences	2,000	100	120
Applied sciences	1,800	90	84
Engineering	2,200	110	91
TOTAL	9,000	450	450

the present example, if there are five faculties we would have five strata, with the numbers in each stratum being one-twentieth of the total for each faculty, as in Table 8.1, which also shows a hypothetical outcome of using a simple random sample, which results in a distribution of students across faculties that does not mirror the population all that well.

The advantage of stratified random sampling in a case like this is clear: it ensures that the resulting sample will be distributed in the same way as the population in terms of the stratifying criterion. If you use a simple random or systematic sampling approach, you *may* end up with a distribution like that of the stratified sample, but it is unlikely. Two points are relevant here. First, you can conduct stratified sampling sensibly only when it is relatively easy to identify and allocate units to strata. If it is not possible or it would be very difficult to do so, stratified sampling will not be feasible. Second, you can use more than one stratifying criterion. Thus, it may be that you would want to stratify by both faculty and gender or

faculty and whether students are undergraduates or postgraduates. If it is feasible to identify students in terms of these stratifying criteria, it is possible to use pairs of criteria or several criteria (such as faculty membership plus gender plus undergraduate/postgraduate).

Stratified sampling is really feasible only when the relevant information is available. In other words, when

data are available that allow the ready identification of members of the population in terms of the stratifying criterion (or criteria), it is sensible to employ this sampling method. But it is unlikely to be economical if the identification of population members for stratification purposes entails a great deal of work because there is no available listing in terms of strata.



Student experience

Probability sampling for a student project

Joe Thompson describes the sampling procedure that he and the other members of his team used for their study of students living in halls of residence at the University of East Anglia as a **stratified random sample**. The following description suggests that they employed a systematic sampling approach for finding students within halls.

Stratified random sampling was used to decide which halls of residence each member of the research team would go to and obtain questionnaire responses. This sampling method was the obvious choice as it meant there could be no fixing/bias to which halls the interviewee would go to and also maintained the representative nature of the research.

The stratified random sampling method known as the 'random walk process' was used when conducting the interviews. Each member of the research group was assigned a number between 4 and 8 as a sampling fraction gap: I was assigned the number 7 and 'Coleman house block 1' as my accommodation block. This meant that, when conducting my interviews, I would go to Coleman house and knock on the 7th door, and then the 14th door, adding 7 each time, until I had completed five interviews. If I encountered a lack of response from the 6th door, I would return to the first flat but add one each time to avoid periodicity. This sampling method was determined by the principles of standardization, reliability, and validity.



To read more about Joe's research experiences, go to the Online Resource Centre that accompanies this book at: www.oxfordtextbooks.co.uk/orc/brymansrm4e/

Multi-stage cluster sampling

In the example we have been dealing with, students to be interviewed are located in a single university. Interviewers will have to arrange their interviews with the sampled students, but, because they are all close together (even in a split-site university), they will not be involved in a lot of travel. However, imagine that we wanted a *national* sample of students. It is likely that interviewers would have to travel the length and breadth of the UK to interview the sampled students. This would add a great deal to the time and cost of doing the research. This kind of problem occurs whenever the aim is to interview a sample that is to be drawn from a widely dispersed population, such as a national population, or a large region, or even a large city.

One way in which it is possible to deal with this potential problem is to employ **cluster sampling**. With cluster

sampling, the primary sampling unit (the first stage of the sampling procedure) is not the units of the population to be sampled but groupings of those units. It is the latter groupings or aggregations of population units that are known as *clusters*. Imagine that we want a nationally representative sample of 5,000 students. Using simple random or systematic sampling would yield a widely dispersed sample, which would result in a great deal of travel for interviewers. One solution might be to sample universities and then students from each of the sampled universities. A probability sampling method would need to be employed at each stage. Thus, we might randomly sample ten universities from the entire population of universities, thus yielding ten clusters, and we would then interview 500 randomly selected students at each of the ten universities.

Now imagine that the result of sampling ten universities gives the following list:

- Glasgow Caledonian
- Edinburgh
- Teesside
- Sheffield
- University College Swansea
- Leeds Metropolitan
- University of Ulster
- University College London
- Southampton
- Loughborough

This list is fine, but interviewers could still be involved in a great deal of travel, since the ten universities are quite a long way from each other. North American and

Australian readers who examine this last comment by looking at a map of the United Kingdom may view the universities as in fact very close to each other!

One solution is likely to be to group all UK universities by standard region (see Research in focus 8.1 for an example of this kind of approach) and randomly to sample two standard regions. Five universities might then be sampled from each of the two lists of universities and then 500 students from each of the ten universities. Thus, there are separate stages:

- group UK universities by standard region and sample two regions;
- sample five universities from each of the two regions;
- sample 500 students from each of the ten universities.



Research in focus 8.1

An example of a multi-stage cluster sample

For their study of social class in modern Britain, Marshall et al. (1988: 288) designed a sample 'to achieve 2,000 interviews with a random selection of men aged 16–64 and women aged 16–59 who were not in full-time education'.

- Sampling parliamentary constituencies
 - Parliamentary constituencies were ordered by standard region (there are eleven).
 - Constituencies were allocated to one of three population density bands within standard regions.
 - These subgroups were then reordered by political party voted to represent the constituency at the previous general election.
 - These subgroups were then listed in ascending order of percentage in owner–occupation.
 - 100 parliamentary constituencies were then sampled.
 - Thus, parliamentary constituencies were stratified in terms of four variables: standard region; population density; political party voted for in last election; and percentage of owner–occupation.
- Sampling polling districts
 - Two polling districts were chosen from each sampled constituency.
- Sampling individuals
 - Nineteen addresses from each sampled polling district were systematically sampled.
 - One person at each address was chosen according to a number of pre-defined rules.

In a sense, cluster sampling is always a multi-stage approach, because one always samples clusters first, and then something else—either further clusters or population units—is sampled.

Many examples of multi-stage cluster sampling entail stratification. We might, for example, want to stratify universities in terms of whether they are 'old' or 'new' universities—that is, those that received their charters after the 1991 White Paper for Higher Education, *Higher*

Education: A New Framework. In each of the two regions, we would group universities along the old/new university criterion and then select two or three universities from each of the two strata per region.

Research in focus 8.1 provides an example of a multi-stage cluster sample. It entailed three stages: the sampling of parliamentary constituencies, the sampling of polling districts, and the sampling of individuals. In a way, there are four stages, because addresses are

sampled from polling districts and then individuals are sampled from each address. However, Marshall et al. (1988) present their sampling strategy as involving just three stages. Parliamentary constituencies were stratified by four criteria: standard region, population density, voting behaviour, and owner–occupation.

The advantage of multi-stage cluster sampling should be clear by now: it allows interviewers to be far more

geographically concentrated than would be the case if a simple random or stratified sample were selected. The advantages of stratification can be capitalized upon because the clusters can be stratified in terms of strata. However, even when a very rigorous sampling strategy is employed, sampling error cannot be avoided, as the example in Research in focus 8.2 shows.



Research in focus 8.2 The 1992 British Crime Survey

The British Crime Survey (BCS) is a regular survey, funded by the Home Office, of a national sample drawn from the populations of England and Wales. The survey was conducted on eight occasions between 1982 and 2000 and has been conducted annually since 2001. In each instance, over 10,000 people have been interviewed. The main object of the survey is to glean information on respondents' experiences of being victims of crime. There is also a self-report component in which a selection of the sample are interviewed on their attitudes to crime and to report on crimes they have committed. Before 1992, the BCS used the electoral register as a sampling frame. Relying on a register of the electorate as a sampling frame is not without problems in spite of appearing robust: it omits any persons who are not registered, a problem that was exacerbated by the Community Charge (poll tax), which resulted in a significant amount of non-registration, as some people sought to avoid detection in order not to have to pay the tax. In 1992 the Postcode Address File was employed as a sampling frame and has been used since then. Its main advantage over the electoral register as a sampling frame is that it is updated more frequently. It is not perfect, because the homeless will not be accessible through it. The BCS sample itself is a stratified multi-stage cluster sample. The sampling procedure produced 13,117 residential addresses. Like most surveys, there was some non-response, with 23.3 per cent of the 13,117 addresses not resulting in a 'valid' interview. Just under half of these cases were the result of an outright refusal. In spite of the fact that the BCS is a rigorously selected and very large sample, an examination of the 1992 survey by Elliott and Ellingworth (1997) shows that there is some sampling error. By comparing the distribution of survey respondents with the 1991 census, they show that certain social groups are somewhat under-represented, most notably: owner–occupiers, households in which no car is owned, and male unemployed. However, Elliott and Ellingworth show that, as the level of property crime in postcode address sectors increases, the response rate (see Key concept 8.2) decreases. In other words, people who live in high-crime areas tend to be less likely to agree to be interviewed. How far this tendency affects the BCS data is difficult to determine, but the significance of this brief example is that, even when a sample of this quality is selected, the existence of sampling and non-sampling error cannot be discounted. The potential for a larger spread of errors when levels of sampling rigour fall short of a sample like that selected for the BCS is, therefore, considerable.



The qualities of a probability sample

The reason why probability sampling is such an important procedure in social survey research is that it is possible to make inferences from information about a random sample to the population from which it was selected. In other words, we can generalize findings

derived from a sample to the population. This is not to say that we treat the population data and the sample data as the same. If we take the example of the level of alcohol consumption in our sample of 450 students, which we will treat as the number of units of alcohol consumed in

the previous seven days, we will know that the **mean** number of units consumed by the sample ($\bar{x}$) can be used to estimate the population mean (μ) but with known margins of error. The mean, or more properly the **arithmetic mean**, is the simple average.

In order to address this point it is necessary to use some basic statistical ideas. These are presented in Tips and skills 'Generalizing from a random sample to the population' and can be skipped if just a broad idea of sampling procedures is required.



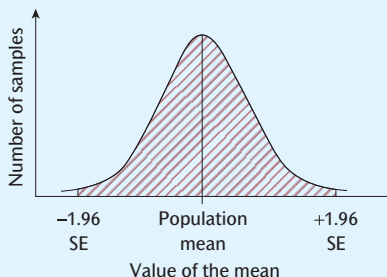
Tips and skills

Generalizing from a random sample to the population

Let us say that the sample mean is 9.7 units of alcohol consumed (the average amount of alcohol consumed in the previous seven days in the sample). A crucial consideration here is: how confident can we be that the mean level of alcohol consumption of 9.7 units is likely to be found in the population, even when probability sampling has been employed? If we take an infinite number of samples from a population, the sample estimates of the mean of the variable under consideration will vary in relation to the population mean. This variation will take the form of a bell-shaped curve known as a *normal distribution* (see Figure 8.8). The shape of the distribution implies that there is a clustering of sample means at or around the population mean. Half the sample means will be at or below the population mean; the other half will be at or above the population mean. As we move to the left (at or lower than the population mean) or the right (at or higher than the population mean), the curve tails off, implying fewer and fewer samples generating means that depart considerably from the population mean. The variation of sample means around the population mean is the *sampling error* and is measured using a statistic known as the **standard error of the mean**. This is an estimate of the amount that a sample mean is likely to differ from the population mean.

Figure 8.8

The distribution of sample means



Notes: 95 per cent of sample means will lie within the shaded area. SE = standard error of the mean.

This consideration is important because sampling theory tells us that 68 per cent of all sample means will lie between ± 1 standard error from the population mean and that 95 per cent of all sample means will lie between ± 1.96 standard errors from the population mean. It is this second calculation that is crucial, because it is at least implicitly employed by survey researchers when they report their statistical findings.

They typically employ 1.96 standard errors as the crucial criterion in how confident they can be in their findings. Essentially, the criterion implies that you can be 95 per cent certain that the population mean lies within + or – 1.96 sampling errors from the sample mean.

If a sample has been selected according to probability sampling principles, we know that we can be 95 per cent certain that the population mean will lie between the sample mean + or – 1.96 multiplied by the standard error of the mean. This is known as the *confidence interval*. If the mean level of alcohol consumption in the previous seven days in our sample of 450 students is 9.7 units and the standard error of the mean is 1.3, we can be 95 per cent certain that the population mean will lie between

$$9.7 + (1.96 \times 1.3)$$

and

$$9.7 - (1.96 \times 1.3)$$

i.e. between 12.248 and 7.152.

If the standard error was smaller, the range of possible values of the population mean would be narrower; if the standard error was larger, the range of possible values of the population mean would be wider.

If a stratified sample is selected, the standard error of the mean will be smaller because the variation between strata is essentially eliminated because the population will be accurately represented in the sample in terms of the stratification criterion or criteria employed. This consideration demonstrates the way in which stratification injects an extra increment of precision into the probability sampling process, since a possible source of sampling error is eliminated.

By contrast, a cluster sample without stratification exhibits a larger standard error of the mean than a comparable simple random sample. This occurs because a possible source of variability between students (i.e. membership of one university rather than another, which may affect levels of alcohol consumption) is disregarded. If, for example, some universities had a culture of heavy drinking in which a large number of students participated, and if these universities were not selected because of the procedure for selecting clusters, an important source of variability would have been omitted. It also implies that the sample mean would be on the low side, but that is another matter.



Sample size

As someone who is known as a teacher of research methods and a writer of books in this area, I often get asked questions about methodological issues. One question that is asked almost more than any other relates to the size of the sample—‘how large should my sample be?’ or ‘is my sample large enough?’ The decision about sample size is not a straightforward one: it depends on a number of considerations, and there is no one definitive answer. This is frequently a source of great disappointment to those who pose such questions. Moreover, most of the time decisions about sample size are affected by considerations of time and cost. Therefore, invariably decisions about sample size represent a compromise between the constraints of time and cost, the need for

precision, and a variety of further considerations that will now be addressed.

Absolute and relative sample size

One of the most basic considerations, and one that is possibly the most surprising, is that, contrary to what you might have expected, it is the *absolute* size of a sample that is important not its *relative* size. This means that a national probability sample of 1,000 individuals in the UK has as much validity as a national probability sample of 1,000 individuals in the USA, even though the latter has a much larger population. It also means that increasing the size of a sample increases the precision of a sample.

This means that the 95 per cent confidence interval referred to in Tips and skills 'Generalizing from a random sample to the population' narrows. However, a large sample cannot *guarantee* precision, so that it is probably better to say that increasing the size of a sample increases the *likely* precision of a sample. This means that, as sample size increases, sampling error decreases. Therefore, an important component of any decision about sample size should be how much sampling error one is prepared to tolerate. The less sampling error one is prepared to tolerate, the larger a sample will need to be. Fowler (1993) warns against a simple acceptance of this criterion. He argues that in practice researchers do not base their

decisions about sample size on a single estimate of a variable. Most survey research is concerned to generate a host of estimates—that is, of the variables that make up the research instrument that is administered. He also observes that it is not normal for survey researchers to be in a position to specify in advance 'a desired level of precision' (Fowler 1993: 34). Moreover, since sampling error will be only one component of any error entailed in an estimate, the notion of using a desired level of precision as a factor in a decision about sample size is not realistic. Instead, to the extent that this notion does enter into decisions about sample size, it usually does so in a general rather than in a calculated way.



Tips and skills

Sample size and probability sampling

As I have said in the text, the issue of sample size is the matter that most often concerns students and others. Basically, this is an area where size really does matter—the bigger the sample, the more representative it is likely to be (provided the sample is randomly selected), regardless of the size of the population from which it is drawn. However, when doing projects, students clearly need to do their research with very limited resources. You should try to find out from your department whether there are any guidelines about whether samples of a minimum size are expected. If there are no such guidelines, you will need to conduct your mini-survey in such a way as to maximize the number of interviews you can manage or the number of postal questionnaires you can send out, given the amount of time and resources available to you. Also, in many if not most cases, a truly random approach to sample selection may not be open to you. The crucial point is to be clear about and to justify what you have done. Explain the difficulties that you would have encountered in generating a random sample. Explain why you really could not include any more in your sample of respondents. But, above all, do not make claims about your sample that are not sustainable. Do not claim that it is representative or that you have a random sample when it is clearly not the case that either of these is true. In other words, be frank about what you have done. People will be much more inclined to accept an awareness of the limits of your sample design than claims about a sample that are patently false. Also, it may be that there are lots of good features about your sample—the range of people included, the good response rate, the high level of cooperation you received from the firm. Make sure you play up these positive features at the same time as being honest about its limitations.

Time and cost

Time and cost considerations become very relevant in this context. In the previous paragraph it is clearly being suggested that the larger the sample size the greater the precision (because the amount of sampling error will be less). However, by and large, up to a sample size of around 1,000, the gains in precision are noticeable as the sample size climbs from low figures of 50, 100, 150, and so on upwards. After a certain point, often in the region of 1,000, the sharp increases in precision become less pronounced, and, although it does not plateau, there is a

slowing-down in the extent to which precision increases (and hence the extent to which the sample error of the mean declines). Considerations of sampling size are likely to be profoundly affected by matters of time and cost at such a juncture, since striving for smaller and smaller increments of precision becomes an increasingly uneconomic proposition. As Hazelrigg (2004: 85) succinctly puts it: 'The larger the size of the sample drawn from a population the more likely ($\bar{x}$) converges to μ ; but the convergence occurs at a decelerating rate (which means that very large samples are decreasingly cost efficient).'

Non-response

However, considerations about sampling error do not end here. The problem of **non-response** should be borne in mind. Most sample surveys attract a certain amount of non-response. Thus, it is likely that only some members of our sample will agree to participate in the research. If it is our aim to ensure as far as possible that 450 students are interviewed and if we think that there may be a 20 per cent rate of non-response, it may be advisable to sample 540–50 individuals, on the grounds that approximately 90 will be non-respondents.

The issue of non-response, and in particular of refusal to participate, is of particular significance, because it has been suggested by some researchers that response rates to social surveys (see Key concept 8.2) are declining in many countries. This implies that there is a growing tendency towards people refusing to participate in social survey research. In 1973 an article in the American magazine *Business Week* carried an article ominously entitled 'The Public Clams up on Survey Takers'. The magazine asked survey companies about their experiences and found considerable concern about declining response rates. Similarly, in Britain, a report from a working party on the Market Research Society's Research and Development Committee in 1975 pointed to similar

concerns among market research companies. However, an analysis of this issue by T. W. Smith (1995) suggests that, contrary to popular belief, there is no consistent evidence of such a decline. Moreover, Smith shows that it is difficult to disentangle general trends in response rates from such variables as the subject matter of the research, the type of respondent, and the level of effort expended on improving the number of respondents to individual surveys. However, an overview of non-response trends in the USA based on non-response rates for various continuous surveys suggests that there is a decline in the preparedness of households to participate in surveys (Groves et al. 2004). Further evidence comes from a study by Baruch (1999) of questionnaire-based articles published in 1975, 1985, and 1995 in five academic journals in the area of management studies. This article found an average (mean) response rate of 55.6 per cent, though with quite a large amount of variation around this average. The average response rate over the three years was 64.4 per cent in 1975, 55.7 per cent in 1985, and 48.4/52.2 per cent in 1995. Two percentages were provided for 1995 because the larger figure includes a journal that publishes a lot of research based on top managers, who tend to produce a poorer response rate. Response rates were found that were as low as 10 per cent and 15 per cent.



Key concept 8.2 What is a response rate?

The notion of a response rate is a common one in social survey research. When a social survey is conducted, whether by structured interview or by self-completion questionnaire, it is invariably the case that some people who are in the sample refuse to participate (referred to as non-response). The response rate is, therefore, the percentage of a sample that does, in fact, agree to participate. However, the calculation of a response rate is a little more complicated than this. First, not everyone who replies will be included: if a large number of questions are not answered by a respondent or if there are clear indications that he or she has not taken the interview or questionnaire seriously, it is better to employ only the number of *usable* interviews or questionnaires as the numerator. Similarly, it also tends to occur that not everyone in a sample turns out to be a suitable or appropriate respondent or can be contacted. Thus the response rate is calculated as follows:

$$\frac{\text{number of usable questionnaires}}{\text{total sample} - \text{unsuitable or uncontactable members of the sample}} \times 100$$

A further interesting issue in connection with non-response is that of how far researchers should go in order to boost their response rates. In Chapter 10, a number of steps that can be taken to improve response rates to postal questionnaires, which are particularly prone to

poor response rates, are discussed. However, boosting response rates to interview-based surveys can prove expensive. Teitler et al. (2003) present a discussion of the steps taken to boost the response rate of a US sample that was hard to reach—namely, both parents of newly

born children, where most of the parents were not married. They found that, although there was evidence that increasing the response rate from an initial 68 per cent to 80 per cent meant that the final sample resembled more closely the population from which the sample had been taken, diminishing returns undoubtedly set in. In other words, the improvements in the characteristics of the sample necessitated a disproportionate outlay of resources. However, this is not to say that steps should not be taken to improve response rates. For example, following up respondents who do not initially respond to a postal questionnaire invariably results in an improved response rate at little additional cost. A study based on a survey of New Zealand residents by Brennan and Charbonneau (2009) provides unequivocal evidence of the improvement in response rate that can be achieved by at least two follow-up mailings to respondents to postal questionnaire surveys, which tend to achieve lower response rates than comparable interview-based surveys. A chocolate sent with the questionnaire helps too apparently!

As the previously mentioned study of response rates by Baruch (1999) suggests, there is wide variation in the response rates that social scientists achieve when they conduct surveys. It is difficult to arrive at clear indications of what is expected from a response rate. Baruch's study focused on research in business organizations,

and, as he notes, when top managers are the focus of a survey, the response rate tends to be noticeably lower. In the survey component of the Cultural Capital and Social Exclusion (CCSE) project referred to in Research in focus 2.9, the initial main sample constituted a 53 per cent response rate (Bennett et al. 2009). The researchers decided to supplement the initial sample in various ways, one of which was to have an ethnic boost sample, in large part because the main sample did not include sufficient numbers of ethnic-minority members. However, the response rate from the ethnic boost sample was substantially below that achieved for the main sample. The researchers write: 'In general, ethnic boosts tend to have lower response rates than cross-sectional surveys' (Thomson 2004: 10). There is a sense, then, that what might be anticipated to be a reasonable response rate varies according to the type of sample and the topics covered by the interview or questionnaire. While it is obviously desirable to do one's best to maximize a response rate, it is also important to be open about the limitations of a low response rate in terms of the likelihood that findings will be biased. In the future, it seems likely that, given that there are likely to be limits on the degree to which a survey researcher can boost a response rate, more and more effort will go into refining ways of estimating and correcting for anticipated biases in findings (Groves 2006).



Research in focus 8.3

The problem of non-response

In December 2006 an article in *The Times* reported that a study of the weight of British children had been hindered because many families declined to participate. The study was commissioned by the Department of Health and found that, for example, among those aged 10 or 11, 14 per cent were overweight and 17 per cent were obese. However, *The Times* writer notes that a report compiled by the Department of Health on the research suggests that such figures are 'likely systematically to underestimate the prevalence of overweight and obesity' (quoted in Hawkes 2006: 24). The reason for this bias in the statistics is that parents were able to refuse to let their children participate, and those whose children were heavier were more likely to do so. As a result, the sample was biased towards those who were less heavy. The authors of the report drew the inference about sampling bias because they noted that more children were recorded as obese in areas where there was a poorer response rate.

Heterogeneity of the population

Yet another consideration is the homogeneity and heterogeneity of the population from which the sample is to be taken. When a population is very heterogeneous, like a whole country or city, a larger sample will be

needed to reflect the varied population. When it is relatively homogeneous, such as a population of students or of members of an occupation, the amount of variation is less and therefore the sample can be smaller. The implication of this is that, the greater the heterogeneity of a population, the larger a sample will need to be.

Kind of analysis

Finally, researchers should bear in mind the *kind of analysis* they intend to undertake. A case in point here is the **contingency table**. A contingency table shows the relationship between two variables in tabular form. It shows how variation in one variable relates to variation in another variable. To understand this point, consider the basic structure of a table in the study by Marshall et al. (1988) of social class in Britain. This research was referred to in Research in focus 8.1. The table is based on the 589 cohabiting couples (1,178 people) of the sample in which both partners are employed in paid work. The authors aim to show in the table how far couples are of the same or a different social class in terms of Goldthorpe's seven-category scheme for classifying social class. The result is a table in which, because each variable comprises 7 categories, there are 49 **cells** in the table (i.e. 7×7). In order for there to be an adequate number of cases in each

cell, a fairly large sample was required. Imagine that Marshall et al. had conducted a survey on a much smaller sample in which they ended up with just 150 couples. If the same kind of analysis as Marshall et al. carried out was conducted, it would be found that these 150 couples would be very dispersed across the 49 cells of the table. It is likely that many of the cells would be empty or would have very small numbers in them, which would make it difficult to make inferences about what the table showed. In fact, quite a lot of the cells in the actual table in Marshall et al. have very small numbers in them (8 cells contain 1 or 0). This problem would have been even more pronounced if they had ended up with a much smaller sample of couples. Consequently, considerations of sample size should be sensitive to the kinds of analysis that will be subsequently required, such as the issue of the number of cells in a table. In a case such as this, a larger sample will be necessitated by the nature of the analysis to be conducted as well as the nature of the variables in question.



Types of non-probability sampling

The term 'non-probability sampling' is essentially an umbrella term to capture all forms of sampling that are not conducted according to the canons of probability sampling outlined above. It is not surprising, therefore, that the term covers a wide range of different types of sampling strategy, at least one of which—the **quota sample**—is claimed by some practitioners to be almost as good as a probability sample. In this section we will cover three main types of non-probability sample: the **convenience sample**; the **snowball sample**; and the **quota sample**.

Convenience sampling

A convenience sample is one that is simply available to the researcher by virtue of its accessibility. Imagine that a researcher who teaches education at a university is interested in the kinds of features that teachers look for in their headmasters. The researcher might administer a questionnaire to several classes of students, all of whom are teachers taking a part-time master's degree in education. The chances are that the researcher will receive all or almost all of the questionnaires back, so that there will be a good response rate. The findings may prove quite interesting, but the problem with such a sampling

strategy is that it is impossible to generalize the findings, because we do not know of what population this sample is representative. They are simply a group of teachers who are available to the researcher. They are almost certainly not representative of teachers as a whole—the very fact they are taking this degree programme marks them off as different from teachers in general.

This is not to suggest that convenience samples should never be used. Let us say that our lecturer/researcher is developing a battery of questions that are designed to measure the leadership preferences of teachers. It is highly desirable to pilot such a research instrument before using it in an investigation, and administering it to a group that is not a part of the main study may be a legitimate way of carrying out some preliminary analysis of such issues as whether respondents tend to answer in identical ways to a question, or whether one question is often omitted when teachers respond to it. In other words, for this kind of purpose, a convenience sample may be acceptable though not ideal. A second kind of context in which it may be at least fairly acceptable to use a convenience sample is when the chance presents itself to gather data from a convenience sample and it represents too good an opportunity to miss. The data will not allow definitive findings to be generated, because of the

problem of generalization, but they could provide a springboard for further research or allow links to be forged with existing findings in an area.

It also perhaps ought to be recognized that convenience sampling probably plays a more prominent role than is sometimes supposed. Certainly, in the field of organization studies it has been noted that convenience samples are very common and indeed are more prominent

than are samples based on probability sampling (Bryman 1989a: 113–14). Social research is also frequently based on convenience sampling. Research in focus 8.4 contains an example of the use of convenience samples in social research. Probability sampling involves a lot of preparation, so that it is frequently avoided because of the difficulty and costs involved.



Research in focus 8.4

A convenience sample

Miller et al. (1998) were interested in theories concerning the role of shopping in relation to the construction of identity in modern society. Since many discussions of this issue have been concerned with shopping centres (malls), they undertook a study that combined quantitative and qualitative research methods in order to explore the views of shoppers at two London shopping centres: Brent Cross and Wood Green. One phase of the research entailed structured interviews with shoppers leaving the centres. The interviews were conducted mainly during weekdays in June and July 1994. Shoppers were chiefly questioned as they left the main exits, though some questioning at minor exits also took place. The authors tell us: 'We did not attempt to secure a quota [see below] or random sample but asked every person who passed by, and who did not obviously look in the other direction or change their path, to complete a questionnaire' (Miller et al. 1998: 55). Such a sampling strategy produces a convenience sample because only people who are visiting the centre and who are therefore self-selected by virtue of their happening to choose to shop at these times can be interviewed.

Snowball sampling

In certain respects, snowball sampling is a form of convenience sample, but it is worth distinguishing because it has attracted quite a lot of attention over the years. With this approach to sampling, the researcher makes initial contact with a small group of people who are relevant to the research topic and then uses these to establish

contacts with others. I used an approach like this to create a sample of British visitors to Disney theme parks (Bryman 1999).

Research in focus 8.5 describes the generation of a snowball sample of marijuana-users for what is often regarded as a classic study of drug use. Becker's comment on this method of creating a snowball sample is interesting: 'The sample is, of course, in no sense



Research in focus 8.5

A snowball sample: Becker's study of marijuana-users

In an article first published in 1953, Becker (1963) reports on how he generated a sample of marijuana-users. He writes:

I conducted fifty interviews with marijuana users. I had been a professional dance musician for some years when I conducted this study and my first interviews were with people I had met in the music business. I asked them to put me in contact with other users who would be willing to discuss their experiences with me. . . . Although in the end half of the fifty interviews were conducted with musicians, the other half covered a wide range of people, including laborers, machinists, and people in the professions. (Becker 1963: 45–6)

“random”; it would not be possible to draw a random sample, since no one knows the nature of the universe from which it would have to be drawn’ (Becker 1963: 46). What Becker is essentially saying here (and the same point applies to my study of Disney theme park visitors) is that there is no accessible sampling frame for the population from which the sample is to be taken and that the difficulty of creating such a sampling frame means that a snowball sampling approach is the only feasible one. Moreover, even if one could create a sampling frame of marijuana-users or of British visitors to Disney theme parks, it would almost certainly be inaccurate straight away, because this is a shifting population. People will constantly be becoming and ceasing to be marijuana-users, while new theme park visitors are arriving all the time.

The problem with snowball sampling is that it is very unlikely that the sample will be representative of the population, though, as I have just suggested, the very notion of a population may be problematic in some circumstances. However, by and large, snowball sampling is used not within a quantitative research strategy, but within a qualitative one: both Becker’s and my study were carried out within a qualitative research framework. Concerns about external validity and the ability to generalize do not loom as large within a qualitative research strategy as they do in a quantitative research one (see Chapter 17). In qualitative research, the orientation to sampling is more likely to be guided by a preference for *theoretical sampling* than with the kind of statistical sampling that has been the focus of this chapter (see Key concept 18.3). There is a much better ‘fit’ between snowball sampling and the theoretical sampling strategy of qualitative research than with the statistical sampling approach of quantitative research. This is not to suggest that snowball sampling is entirely irrelevant to quantitative research: when the researcher needs to focus upon or to reflect relationships between people, tracing connections through snowball sampling may be a better approach than conventional probability sampling (Coleman 1958).

Quota sampling

Quota sampling is comparatively rarely employed in academic social research, but is used intensively in commercial research, such as market research and political opinion polling. The aim of quota sampling is to produce a sample that reflects a population in terms of the relative proportions of people in different categories, such as gender, ethnicity, age groups, socio-economic groups,

and region of residence, and in combinations of these categories. However, unlike a stratified sample, the sampling of individuals is not carried out randomly, since the final selection of people is left to the interviewer. Information about the stratification of the UK population or about certain regions can be obtained from sources like the census and from surveys based on probability samples such as the General Household Survey, British Social Attitudes, and the British Household Panel Survey.

Once the categories and the number of people to be interviewed within each category (known as *quotas*) have been decided upon, it is then the job of interviewers to select people who fit these categories. The quotas will typically be interrelated. In a manner similar to stratified sampling, the population may be divided into strata in terms of, for example, gender, social class, age, and ethnicity. Census data might be used to identify the number of people who should be in each subgroup. The numbers to be interviewed in each subgroup will reflect the population. Each interviewer will probably seek out individuals who fit several subgroup quotas. Accordingly, an interviewer may know that among the various subgroups of people he or she must find, and interview, five Asian, 25–34-year-old, lower-middle-class females in the area in which the interviewer has been asked to work (say, the Wirral). The interviewer usually asks people who are available to him or her about their characteristics (though gender will presumably be self-evident) in order to determine their suitability for a particular subgroup. Once a subgroup quota (or a combination of subgroup quotas) has been achieved, the interviewer will no longer be concerned to locate individuals for that subgroup.

The choice of respondents is left to the interviewer, subject to the requirement of all quotas being filled, usually within a certain time period. Those of you who have ever been approached on the street by a person toting a clipboard and interview schedule and have been asked about your age, occupation, and so on, before being asked a series of questions about a product or whatever, have almost certainly encountered an interviewer with a quota sample to fill. Sometimes, he or she will decide not to interview you because you do not meet the criteria required to fill a quota. This may be due to a quota already having been filled or to the criteria for exclusion meaning that a person with a certain characteristic you possess is not required.

A number of criticisms are frequently levelled at quota samples.

- Because the choice of respondent is left to the interviewer, the proponents of probability sampling

argue that a quota sample cannot be representative. It may accurately reflect the population in terms of superficial characteristics, as defined by the quotas. However, in their choice of people to approach, interviewers may be unduly influenced by their perceptions of how friendly people are or by whether the people make eye contact with the interviewer (unlike most of us, who look at the ground and shuffle past as quickly as possible because we do not want to be bothered in our leisure time).

- People who are in an interviewer's vicinity at the times he or she conducts interviews, and are therefore available to be approached, may not be typical. There is a risk, for example, that people in full-time paid work may be under-represented and that those who are included in the sample are not typical.
- The interviewer is likely to make judgements about certain characteristics in deciding whether to approach a person, in particular, judgements about age. Those judgements will sometimes be incorrect—for example, when someone who is eligible to be interviewed, because a quota that he or she fits is unfilled, is not approached because the interviewer makes an incorrect judgement (for example, that the person is older than he or she looks). In such a case, a possible element of bias is being introduced.
- It has also been argued that the widespread use of social class as a quota control can introduce difficulties, because of the problem of ensuring that interviewees are properly assigned to class groupings (Moser and Kalton 1971).
- It is not permissible to calculate a standard error of the mean from a quota sample, because the non-random method of selection makes it impossible to calculate the range of possible values of a population.
- Because calling back is not required, a quota sample is easier to manage. It is not necessary to keep track of people who need to be recontacted or to keep track of refusals. Refusals occur, of course, but it is not necessary (and indeed it is not possible) to keep a record of which respondents declined to participate.
- When speed is of the essence, a quota sample is invaluable when compared to the more cumbersome probability sample. Newspapers frequently need to know how a national sample of voters feel about a certain topic or how they intend to vote at that time. Alternatively, if there is a sudden major news event, such as a terrorist incident like the London bombs of July 2005, the news media may seek a more or less instant picture of the nation's views about personal security or people's responses more generally. Again, a quota sample will be much faster.
- As with convenience sampling, it is useful for conducting development work on new measures or on research instruments. It can also be usefully employed in relation to exploratory work from which new theoretical ideas might be generated.
- Although the standard error of the mean should not be computed for a quota sample, it frequently is. As Moser and Kalton (1971) observe, some writers argue that the use of a non-random method in quota sampling should not act as a barrier to such a computation because its significance as a source of error is small when compared to other errors that may arise in surveys (see Figure 8.9). However, they go on to argue that at least with random sampling the researcher can calculate the amount of sampling error and does not have to be concerned about its potential impact.

All this makes the quota sample look a poor bet, and there is no doubt that it is not favoured by academic social researchers. It does have some arguments in its favour, however.

- It is undoubtedly cheaper and quicker than an interview survey on a comparable probability sample. For example, interviewers do not have to spend a lot of time travelling between interviews.
- Interviewers do not have to keep calling back on people who were not available at the time they were first approached.

There is some evidence to suggest that, when compared to random samples, quota samples often result in biases. They under-represent people in lower social strata, people who work in the private sector and manufacturing, and people at the extremes of income, and they over-represent women in households with children and people from larger households. On the other hand, it has to be acknowledged that probability samples are often biased too—for example, it is often suggested that they under-represent men and those in employment (Marsh and Scarbrough 1990; Butcher 1994).



Limits to generalization

One point that is often not fully appreciated is that, even when a sample has been selected using probability sampling, any findings can be generalized only to the population from which that sample was taken. This is an obvious point, but it is easy to think that findings from a study have some kind of broader applicability. If we take our imaginary study of alcohol consumption among students at a university, any findings could be generalized only to that university. In other words, you should be very cautious about generalizing to students at other universities. There are many factors that may imply that the level of alcohol consumption is higher (or lower) than among university students as a whole. There may be a higher (or lower) concentration of pubs in the university's vicinity, there may be more (or fewer) bars on the campus, there may be more (or less) of a culture of drinking at this university, or the university may recruit a higher (or lower) proportion of students with disposable income. There may be many other factors too. Similarly, we should be cautious of overgeneralizing in terms of locality. Lunt and Livingstone's (1992: 173) study of consumption habits was based on a postal questionnaire sent to '241 people living in or around Oxford during September 1989'. While the authors' findings represent a fascinating insight into modern consumption patterns, we should be cautious about assuming that they can be generalized beyond the confines of Oxford and its environs.

There could even be a further limit to generalization that is implied by the Lunt and Livingstone (1992) sample. They write that the research was conducted in September 1989. One issue that is rarely discussed in this context and that is almost impossible to assess is whether there is a time limit on the findings that are generated. Quite aside from the fact that we need to appreciate that the findings cannot (or at least should not) be generalized beyond the Oxford area, is there a point at which we have to say, 'well, those findings applied to the Oxford area then but things have changed and we can no longer assume that they apply to that or any other locality'? We are, after all, used to thinking that things have changed when there has been some kind of prominent change. To take a simple example: no one would be prepared to assume that the findings of a study in 1980 of university students' budgeting and personal finance habits would apply to students in the early twenty-first century. Quite aside from changes that might have occurred naturally, the erosion and virtual dismantling of the student grant system has changed the ways students finance their education, including perhaps a greater reliance on part-time work (Lucas 1997), a greater reliance on parents, and the use of loans. But, even when there is no definable or recognizable source of relevant change of this kind, there is none the less the possibility (or even likelihood) that findings are temporally specific. Such an issue is impossible to resolve without further research (Bryman 1989b).



Error in survey research

We can think of 'error', a term that has been employed on a number of occasions, as being made up of four main factors (Figure 8.9).

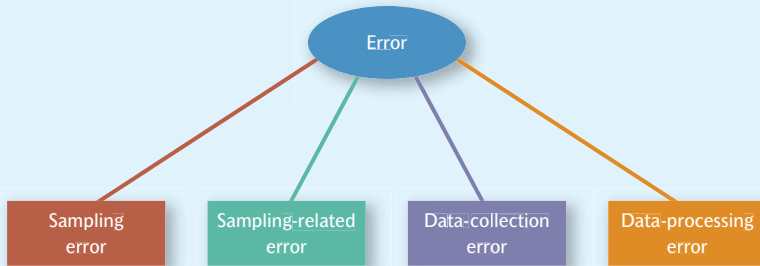
1. **Sampling error.** See Key concept 8.1 for a definition. This kind of error arises because it is extremely unlikely that one will end up with a truly representative sample, even when probability sampling is employed.
2. We can distinguish what might be thought of as *sampling-related error*. This is error that is subsumed under the category *non-sampling error* (see Key concept 8.1) but that arises from activities or events that are related to the sampling process and that are connected with the issue of generalizability or external validity of findings. Examples are an inaccurate sampling frame and non-response.

There is also error that is connected with the implementation of the research process. We might call this *data-collection error*. This source of error includes such factors as: poor question wording in self-completion questionnaires or structured interviews; poor interviewing techniques; and flaws in the administration of research instruments.

3. There is also error that is connected with the implementation of the research process. We might call this *data-collection error*. This source of error includes such factors as: poor question wording in self-completion questionnaires or structured interviews; poor interviewing techniques; and flaws in the administration of research instruments.

Figure 8.9

Four sources of error in social survey research



4. Finally, there is *data-processing error*. This arises from faulty management of data, in particular, errors in the coding of answers.

The third and fourth sources of error relate to factors that are not associated with sampling and instead relate

much more closely to concerns about the validity of measurement, which was addressed in Chapter 7. However, the kinds of steps that need to be taken to keep these sources of error to a minimum in the context of social survey research will be addressed in Chapters 9–11.



Key points

- Probability sampling is a mechanism for reducing bias in the selection of samples.
- Ensure you become familiar with key technical terms in the literature on sampling such as: representative sample; random sample; non-response; population; sampling error; etc.
- Randomly selected samples are important because they permit generalizations to the population and because they have certain known qualities.
- Sampling error decreases as sample size increases.
- Quota samples can provide reasonable alternatives to random samples, but they suffer from some deficiencies.
- Convenience samples may provide interesting data, but it is crucial to be aware of their limitations in terms of generalizability.
- Sampling and sampling-related error are just two sources of error in social survey research.



Questions for review

- What do each of the following terms mean: population; probability sampling; non-probability sampling; sampling frame; representative sample; and sampling and non-sampling error?
- What are the goals of sampling?
- What are the main areas of potential bias in sampling?

Sampling error

- What is the significance of sampling error for achieving a representative sample?

Types of probability sample

- What is probability sampling and why is it important?
- What are the main types of probability sample?
- How far does a stratified random sample offer greater precision than a simple random or systematic sample?
- If you were conducting an interview survey of around 500 people in Manchester, what type of probability sample would you choose and why?
- A researcher positions herself on a street corner and asks 1 person in 5 who walks by to be interviewed. She continues doing this until she has a sample of 250. How likely is she to achieve a representative sample?

The qualities of a probability sample

- A researcher is interested in levels of job satisfaction among manual workers in a firm that is undergoing change. The firm has 1,200 manual workers. The researcher selects a simple random sample of 10 per cent of the population. He measures job satisfaction on a Likert scale comprising ten items. A high level of satisfaction is scored 5 and a low level is scored 1. The mean job satisfaction score is 34.3. The standard error of the mean is 8.58. What is the 95 per cent confidence interval?

Sample size

- What factors would you take into account in deciding how large your sample should be when devising a probability sample?
- What is non-response and why is it important to the question of whether you will end up with a representative sample?

Types of non-probability sample

- Are non-probability samples useless?
- In what circumstances might you employ snowball sampling?
- 'Quota samples are not true random samples, but in terms of generating a representative sample there is little difference between them, and this accounts for their widespread use in market research and opinion polling.' Discuss.

Limits to generalization

- 'The problem of generalization to a population is not just to do with the matter of getting a representative sample.' Discuss.

Error in survey research

- 'Non-sampling error, as its name implies, is concerned with sources of error that are not part of the sampling process.' Discuss.

**Online Resource Centre**

www.oxfordtextbooks.co.uk/orc/brymansrm4e/

Visit the Online Resource Centre that accompanies this book to enrich your understanding of sampling. Consult web links, test yourself using multiple choice questions, and gain further guidance and inspiration from the Student Researcher's Toolkit.